Science of the Total Environment 951 (2024) 175740

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Quantitative fecal pollution assessment with bacterial, viral, and molecular methods in small stream tributaries

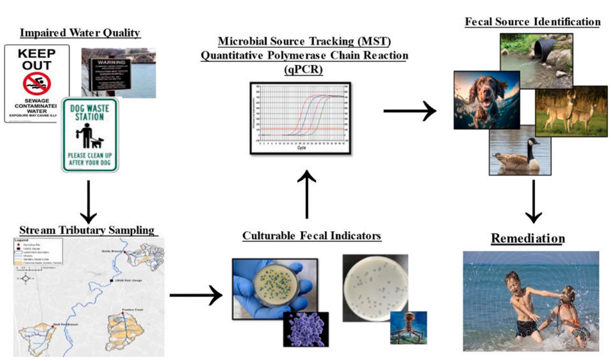
Brian R. McMinn [a,*], Asja Korajkic [a], Julie Kelleher [a], Adam Diedrich [a], Adin Pemberton [a], Jessica R. Willis [a], Mano Sivaganesan [a], Brooke Shireman [b], Andrew Doyle [b], Orin C. Shanks [a]

[a] *Center for Environmental Measurement and Modeling, Office of Research and Development, United States Environmental Protection Laboratory, 26 West Martin Luther King Drive, Cincinnati, OH 45268, United States*
[b] *Sanitation District No. 1 of Northern Kentucky, 1045 Eaton Drive, Fort Wright, KY 41017, United States*

HIGHLIGHTS

- Stream catchments were selected based on land use scenarios impacting water quality.
- Somatic coliphage correlated with *E. coli* ($P < 0.0001$) and enterococci ($P = 0.0002$).
- Somatic coliphage exceeded the 600 PFU/100mL threshold in 36% of samples tested.
- Catchment land use was linked with specific fecal sources impacting water quality.
- Precipitation drove human and ruminant fecal marker occurrence in the catchments

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:
Coliphage
Microbial source tracking
Water quality
Fecal indicator bacteria
Stream tributary

ABSTRACT

Stream water quality can be impacted by a myriad of fecal pollution sources and waste management practices. Identifying origins of fecal contamination can be challenging, especially in high order streams where water samples are influenced by pollution from large drainage areas. Strategic monitoring of tributaries can be an effective strategy to identify conditions that influence local water quality. Water quality is assessed using fecal indicator bacteria (FIB); however, FIB cannot differentiate sources of fecal contamination nor indicate the presence of disease-causing viruses. Under different land use scenarios, three small stream catchments were investigated under 'wet' and 'dry' conditions (Scenario 1: heavy residential; Scenario 2: rural residential; and Scenario 3: undeveloped/agricultural). To identify fecal pollution trends, host-associated genetic targets HF183/BacR287 (human), Rum2Bac (ruminant), GFD (avian), and DG3 (canine) were analyzed along with FIB (*Escherichia coli* and enterococci), viral indicators (somatic and F+ coliphage), six general water quality parameters, and local rainfall. Levels of *E. coli* exceeded single sample maximum limits (235 CFU/100 mL) in 70.7 % of samples, enterococci (70 CFU/100 mL) in 100 % of samples, and somatic coliphage exceeded advisory thresholds (600 PFU/L) in 34.1 % of samples. The detection frequency for the human-associated genetic marker was highest in Scenario 3 (50 % of samples) followed by Scenario 2 (46 %), while the ruminant-associated marker was most prevalent in Scenario 1 (64 %). Due to the high proportion of qPCR-based measurements

* Corresponding author.
E-mail address: mcminn.brian@epa.gov (B.R. McMinn).

https://doi.org/10.1016/j.scitotenv.2024.175740
Received 7 May 2024; Received in revised form 21 August 2024; Accepted 21 August 2024
Available online 23 August 2024

below the limit of quantification, a Bayesian data analysis approach was applied to investigate links between host-associated genetic marker occurrence with that of rainfall and fecal indicator levels. Multiple trends associated with small stream monitoring were revealed, emphasizing the role of rainfall, the utility of fecal source information to improve water quality management. And furthermore, water quality monitoring with bacterial or viral methodologies can alter the interpretation of fecal pollution sources in impaired waters.

1. Introduction

Anthropogenic impacts such as stormwater discharges, agricultural activities, and human waste management practices (centralized or decentralized) can be key sources of fecal contamination in surface waters, leading to poor water quality and potential public health risks (Lobos et al., 2024; Mallin et al., 2009; Wittman et al., 2013). In addition, there is a growing body of evidence suggesting that wildlife and companion animal waste can influence water quality in both agricultural and urban scenarios (Field and Samadpour, 2007; Barnes et al., 2018; Garfield and Walker, 2008). Flowing water systems (e.g., rivers and streams) are particularly susceptible to multiple fecal pollution source impacts where the lack of waste management practices combined with domestic and wild animal sources upstream of a sampling site can contribute to water impairment (Cyterski et al., 2022; Vogel et al., 2007; Brooks et al., 2020). Linking a particular fecal pollution source to a specific waste management activity in flowing water scenarios becomes even more challenging in high order systems where water samples are indicative of potential fecal pollution inputs from large drainage areas influenced by numerous tributary catchments. Alternatively, monitoring small streams with reduced drainage areas can be an effective strategy to identify key land use practices that are influencing local water quality (Reynolds et al., 2021; Tarek et al., 2023). Because tributaries typically drain smaller catchments, these systems can be highly sensitive to municipal stormwater discharges and rain events, rapidly increasing flow rates and potentially introducing high concentrations of fecal pollution into flowing water systems (Kistemann et al., 2002; Baral et al., 2018). These observations highlight the need for further development of small stream tributary water quality monitoring strategies to protect public health and enhance water quality management (Brooks et al., 2020).

For over a century, microbial water quality has been assessed using fecal indicator bacteria (FIB) such as *Escherichia coli* and enterococci (Holcomb and Stewart, 2020). These commensal bacteria are typically found in the digestive tracts of warm-blooded animals and their presence in water can signify the presence of fecal contamination. Despite a long history of use, there are many criticisms for using FIB, mainly the lack of evidence suggesting their presence and persistence in environmental waters is indicative of pathogenic enteric viruses (Korajkic et al., 2022; McMinn et al., 2017a). In response, viruses such as coliphage have been proposed as alternative water quality indicators (McMinn et al., 2017a; United States Environmental Protection Agency, 2015). Coliphages are commensal within human and non-human gastrointestinal tracts and share key morphological and structural characteristics to those of enteric viral pathogens (McMinn et al., 2017a). These advantages and others have led to the use of coliphage as a water quality assessment tool by various management institutions worldwide (United States Environmental Protection Agency, 2015; National Health and Medical Research Council, 2008). However, the majority of coliphage research is focused on coastal, marine, and lake environments (Cyterski et al., 2022; Dean and Mitchell, 2022; Zimmer-Faust et al., 2023) with limited information available on occurrence in flowing tributary waters, especially small stream systems.

Fecal indicator FIB-based (bacterial or viral) water quality assessments are also unable to reveal fecal source information, hampering the ability for managers to implement focused mitigation practices to remediate impaired waters impacted by multiple sources. Source information is important because certain contaminants can pose greater risks to human health than others (Field and Samadpour, 2007; Staley et al., 2012). For instance, human fecal polluted waters have a higher likelihood of containing human-specific enteric pathogens, where animal fecal waste can serve as reservoirs for general enteric pathogens. Additionally, identifying sources of fecal pollution present is important as it provides insight into the origins of fecal contamination to ensure appropriate regulatory and water quality remediation efforts can be implemented, especially in areas where more than one fecal pollution source may be present (Harwood et al., 2014). Numerous microbial source tracking (MST) methodologies are available, allowing for the identification of key fecal sources using technologies such as quantitative real-time PCR (qPCR) (Harwood et al., 2014; Green et al., 2019). In general, qPCR methods target host-associated gene sequences that can be quantified and compared to other water quality metrics such as FIB and coliphage as well as rainfall and land use data to enhance water quality management. While there are numerous studies implementing MST methods at coastal beaches, lakes, and rivers (Cyterski et al., 2022; McMinn et al., 2017b; Gitter et al., 2023), there is limited information on the co-occurrence of coliphage and MST host-associated genetic markers in small stream tributary scenarios. Of the studies, experimental designs using inadequate sample volumes ($\leq$ 100 mL) limited the ability to observe measurable levels of viral indicators present (Vogel et al., 2007), or only a single host-associated genetic marker was measured (Price et al., 2023), making it difficult to gain useful insight into the paired measurement data collected.

Here, paired measurements of coliphage, FIB, and multiple MST host-associated genetic markers were used to assess fecal pollution trends in three flowing small stream tributaries under ‘wet’ and ‘dry’ conditions. Sites were selected to represent typical land use scenarios associated with small stream tributaries including residential with public wastewater sewer lines (Scenario 1), rural, residential with septic system wastewater management (Scenario 2), and predominately agriculture and forested area (Scenario 3). Specifically objectives include 1) assess trends in the presence and amount of fecal pollution using both viral (somatic and F+ coliphage) and bacterial (*E. coli* and enterococci) fecal indicator measurements, 2) identify fecal pollution sources with a panel of four host-associated genetic markers indicative of human (HF183/BacR287) (United States Environmental Protection Agency, 2019), dog (DG3) (Green et al., 2014a), avian (GFD) (Li et al., 2021), and ruminant (Rum2Bac) (Raith et al., 2013) animal groups, and 3) apply a Bayesian censored data analysis approach (Cao et al., 2018) to compare host-associated genetic marker concentrations for each catchment scenario and to explore potential links between rainfall, fecal indicator levels (bacterial and viral), and fecal source occurrence. The authors hypothesize that 1) genetic marker occurrence in each catchment will be driven by precipitation events and will align with landuse characteristics in the area, and 2) the use of a novel Bayesian censored data approach will reveal data trends that otherwise would not be apparent regarding the targets (FIB, coliphage, and genetic markers) analyzed.

2. Materials and methods

2.1. Site selection

The Banklick Creek watershed harbors a 30.9 km creek with a 93.7 km^2 catchment situated in Boone and Kenton Counties in Northern Kentucky with six major tributaries. Upwards of 47 % of the Banklick

Creek watershed is developed, with development concentrated in the central and northern (downstream) portions of the watershed (Fig. 1). The headwaters of Banklick Creek are still primarily undeveloped and agricultural in nature. Catchment boundaries and land use information (Table S1) were determined with ArcGIS mapping software (version 10.3; Light Gray Base Map (ESRI), Redlands, CA) using stream and elevation data layers from the National Hydrography Dataset for Kentucky (Systems, H, 2012) and the 2020 County Contours data sets (*LinkGIS Topo_2020_CC_KC Geospatial Data Presentation Form: vector digital data,* 2021) respectively, and then combined with local land use data such as property valuation assessments and SD1 billing account information (Sanitation District No.1 of Northern Kentucky, 2023). Sampling sites were strategically located at the catchment outlet point where a respective headwater stream intersects a catchment boundary (Fig. 1). Based on upstream catchment land use activities, three small stream tributaries representing different land use scenarios were selected within the Banklick Creek watershed. For Scenario 1, a heavily residential catchment (Holds Branch; HB), contained 1298 residential parcels, 6241 m/km^2 of public sewer lines, and 57 municipal separate storm sewer system (MS4) outfalls. Scenario 2 (Fowler Creek; FC) is a rural residential catchment containing a high proportion of septic systems (444 of 791 residential parcels), 64 MS4 outfalls, and the highest total land area (8.96 km^2) of the three catchments. For Scenario 3 (Wolf Creek; WC), the least developed catchment, contained the highest agricultural land area (3.04 km^2), while having the least residential parcels (n = 79), public sewer lines (792 m/km^2), septic systems (n = 70), and MS4 outfalls (n = 0) of the three catchments.

2.2. Precipitation

To monitor precipitation, a reference United States Geological Survey (USGS) gauge was selected based on proximity to catchments (USGS 03254550, Highway 1829 Erlanger, KY) where hourly measurements over the study period could be collected (Fig. 1). Due to variable land use activities in selected catchments, a broad definition for a 'wet' sample event was used (> 0 mm cumulative precipitation 48 h prior to sampling).

2.3. Sample collection

A total of 41 water samples were collected across 14 weekly sampling events between June and September 2022. All samples were collected at the approximate depth of 25 cm, in the morning at the same time each day. At each site, water temperature (°C), conductivity (μS/cm), pH, and turbidity (FNU) were measured using a YSI ProDDS water quality meter (Yellow Springs Instruments Inc., Yellow Springs, Ohio) according to manufacturer's recommendations. Total nitrogen (mg/L) and phosphorus (mg/L) were measured on site using a handheld Hach Colorimeter II (Hach, Loveland, CO) according to manufacturer's specifications. While methods employing colorimetry represent a less accurate means of assessing chemical concentrations in water than those using ion chromatography technologies (EPA Method 300.1), results have been shown to be comparable between the methodologies (Bianchi et al., 1995; Bergquist et al., 2016; United States Environmental Protection Agency, 1993). On each sampling date, a 3 L grab sample was obtained for each sampling event, using sterile polypropylene bottles. Samples were returned to the laboratory on ice for processing the same day (holding time < 1 h).

2.4. Cultivation sample analyses

2.4.1. Fecal indicator bacteria enumeration

Water samples (between 1 mL and 25 mL) were processed in replicates (n = 3), though pre-wetted [sterile 1× phosphate buffered saline (PBS)] 47 mm cellulose acetate filters with a nominal pore size of 0.45 μm (Thermo Fisher Scientific, Grand Island, NY) and three negative controls consisting of 10 mL of 1× PBS (substituted for water sample) were run during each sampling event. Samples for *E. coli* were processed

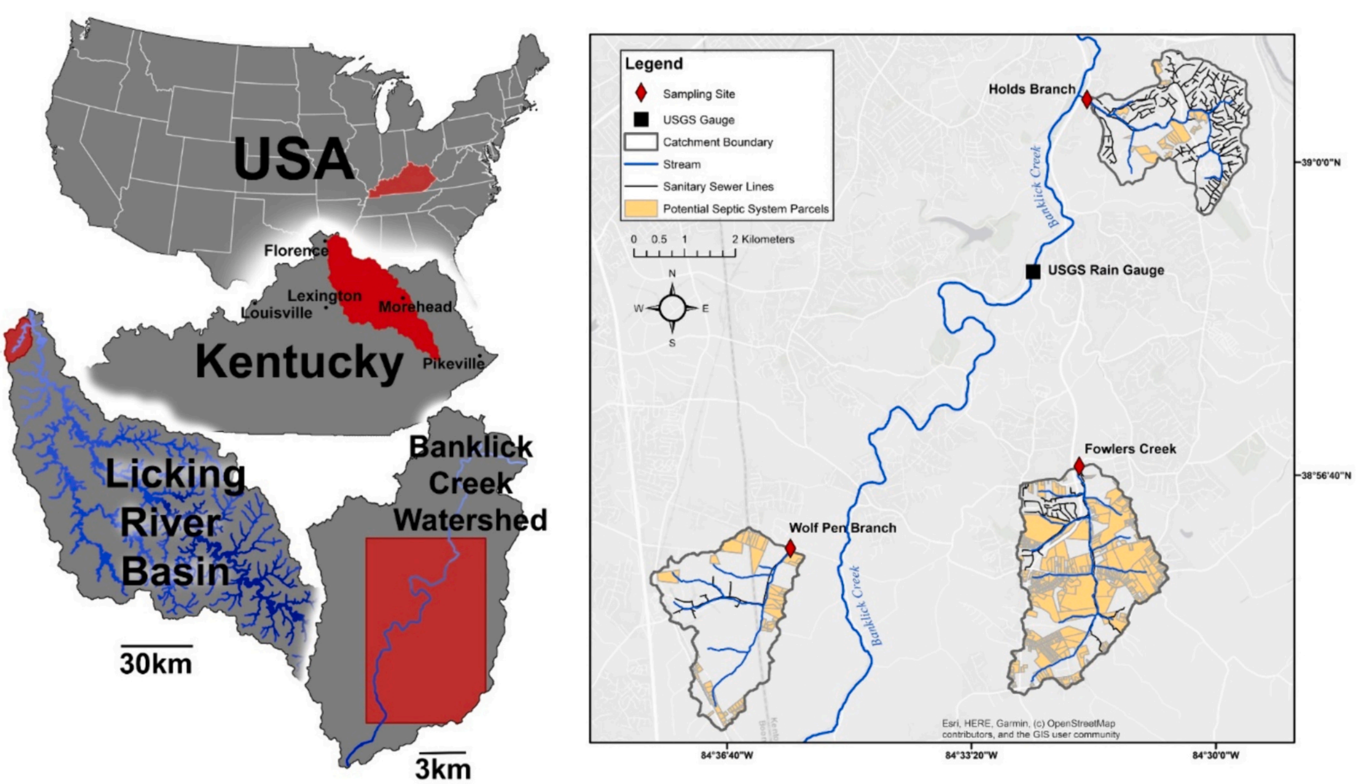


Fig. 1. Geographic information system (GIS) map showing size and location of three catchments included in study. The base layer map is from ESRI (Light Gray Base Map) (*Light Gray Canvas Base Map,* 2023).

according to Environmental Protection Agency (EPA) Method 1603 (United States Environmental Protection Agency, 2014) using a modified mTEC agar (Difco, Sparks, MD). Plates were incubated at 35 °C for 2 h, followed by overnight incubation at 44.5 °C. Red or magenta-colored colonies signifying *E. coli* presence were enumerated and expressed as colony forming units per 100 mL (CFU/100 mL). Samples for enterococci were processed according to EPA Method 1600 (United States Environmental Protection Agency, 2009), using mEI agar (Difco), incubated overnight at 41 °C, and all resulting colonies with blue halos indicative of enterococci growth were recorded and data were expressed as CFU/100 mL.

2.4.2. Coliphage enumeration

Samples from each site were filtered in the laboratory using a previously described dead-end hollowfiber ultrafiltration and single agar layer (D-HFUF-SAL) method (McMinn et al., 2017b). Briefly, 2 L of water was passed through a single use 15S Asahi Kasei Rexeed ultrafilter (Dial Medical Supply, Chester Springs, PA) using a peristaltic pump. Each filter was eluted using 200 mL of an elution solution (0.01 % Tween 80, 0.01 % sodium hexametaphosphate, 0.001 % Antifoam Y-30) (Sigma-Aldrich, St. Louis, MO). An elution solution was circulated through the filter in a clockwise (1 m), counterclockwise (1 m) and finally a clockwise (1 m) pattern to remove attached coliphages from filter surfaces. The resulting eluate (~ 200 mL) was divided evenly (100 mL for somatic and 100 mL for F+ coliphage) and processed using the Single Agar Layer (SAL) procedure, as previously described (McMinn et al., 2017b). Briefly, the filter eluate (100 mL) was mixed with equal parts of a 2× concentration of molten tryptic soy agar containing appropriate concentrations of log-phase host bacteria, antibiotics (nalidixic acid for host ΦX174 or streptomycin/ampicillin for host MS2) and $MgCl_2$ and spread over five 150 mm plates. Negative control plates consisting of agar only (no sample or bacterial host) and agar and host only (no sample) were run in triplicate during each sampling event. Agar plates were incubated at 37 °C for 16–18 h, resulting coliphage plaques were enumerated the following day, and data were adjusted and expressed as plaque forming units per liter (PFU/L).

2.5. Molecular sample analysis

2.5.1. Reference nucleic acid preparations

Reference DNA plasmid constructs consisted of an internal amplification control (IAC; Integrated DNA Technologies, Coralville, Iowa) and the National Institute of Standards and Technology Standard Reference Material® 2917 (SRM 2917; Rockville, Maryland) for qPCR calibration model generation (Willis et al., 2022; Kline et al., 2009). SRM 2917 included five dilutions: Level 1 (10.3 copies/2 μL), Level 2 ($1.11 \cdot 10^2$ copies/2 μL), Level 3 ($1.06 \cdot 10^3$ copies/2 μL), Level 4 ($1.06 \cdot 10^4$ copies/2 μL), and Level 5 ($1.04 \cdot 10^5$ copies/2 μL). The IAC plasmid was linearized by *Sca*I-HF restriction digest (New England BioLabs, Beverly, MA), purified via QIAquick PCR Purification Kit (Qiagen, Valencia, CA), quantified with a Qubit dsDNA HS assay kit on a Qubit 3 Fluorometer (Thermo Fisher Scientific), and diluted in 10 mM Tris 0.5 mM EDTA (pH 9.0) to generate 10^2 copies/2 μL. Salmon sperm DNA (10 μg/mL) was prepared by diluting a commercially available 250 mg stock of DNA sodium salt from salmon testes (Sigma-Aldrich, St. Louis, MO) in 10 mM Tris 0.5 mM EDTA (pH 9.0). All reference DNA materials were stored in low-adhesion microtubes either at 4 °C (SRM 2917 and salmon sperm DNA) or − 20 °C (IAC).

2.5.2. Water filtration and DNADNA extraction

A total of 41 water samples were filtered for DNA extraction. For each sample, 100 mL was filtered through a 0.4 μm polycarbonate filter (Fisher Scientific, Pittsburg, PA). Filters were placed in sterile 2 mL screw cap tubes containing silica bead mill matrix (GeneRite, North Brunswick, JN) and stored at −80 °C (< 6 months) until DNA extraction. Bead milling was achieved with a MP FastPrep-24 (MP BioMedicals, LLC Solon, OH) at 6.0 m/s for 30 s. DNA extraction was performed with the DNA-EZ RW02 kit (GeneRite) according to manufacturer's instructions. DNA was eluted with 100 μL elution buffer into low-adhesion microcentrifuge tubes. Three method extraction blanks (MEB), with molecular grade water substituted for water sample, were performed for each sample processing batch (12 samples/batch). Purified DNA was stored in low adhesion microtubes and stored at 4 °C prior to qPCR amplification (< 24 h).

2.5.3. qPCR amplification

Four host-associated qPCR assays indicative of human (HF183/BacR287), ruminant (Rum2Bac), dog (DG3), and avian (GFD) fecal sources (Green et al., 2014a) were used for MST testing (Table 1). The Sketa22 qPCR assay was included as a sample processing control (SPC) (Haugland et al., 2010). All reactions contained 1× TaqMan Environmental Master Mix (version 2.0; Thermo Fisher Scientific), 0.2 mg/mL bovine serum albumin (Sigma-Aldrich), 1 μM each primer, 80 nM 6-carboxyfluoroscein (FAM)-labeled probe, and 80 nM VIC-labeled probe (HF183/BacR287 only). All reactions contained either 2 μL of DNA extract, SRM 2917, or laboratory grade water in a total reaction volume of 25 μL. HF183/BacR287 duplex reactions also contained 10^2 copies of IAC template. Triplicate reactions were performed for SRM 2917 and Sketa22 testing while all MST experiments included six replicate reactions. Amplifications were conducted on a QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific) with the following thermal cycle profile: 10 min at 95 °C followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C (HF183/BacR287, DG3, Rum2Bac, and Sketa22) and 2 min at 50 °C, 10 min at 95 °C followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C for GFD. The threshold was manually set to 0.03 for all assays. Six no-template controls (NTC) with molecular grade water substituted for template DNA were included with each instrument run to monitor for potential extraneous DNA contamination. Quantification cycle (C_q) values were exported for additional analysis.

2.5.4. qPCR data acceptance metrics

All MST qPCR assays were subject to calibration model acceptance criteria for linearity ($R^2 \geq 0.980$) and amplification efficiency [0.90 to 1.10, $E = 10^{(-1/\text{slope})}$ -1] (Bustin, 2006; LifeTechnologies, 2014). The HF183/BacR287 duplex IAC protocol was used to monitor for amplification inhibition in each water sample (United States Environmental Protection Agency, 2019). Any DNA extract with evidence of amplification inhibition was discarded from the study. Instrument run-specific IAC proficiency testing (HF183/BacR287 NTC VIC C_q standard deviation ≤1.16) was employed to confirm reliable application of amplification inhibition testing (Shanks et al., 2016). A SPC protocol was implemented to monitor for suitable DNA recovery for each water sample and to ensure consistent DNA recovery from one sample batch to another. Samples with unacceptable DNA recovery were excluded from the study based on batch-specific criteria derived from repeated MEB spike recovery experiments. SPC proficiency was determined for each sample batch requiring a standard deviation of Sketa22 qPCR MEB repeated measures of ≤0.62 C_q (Shanks et al., 2016).

2.6. Statistical analyses

A Bayesian Markov Chain Monte Carlo ‘mixed’ calibration model approach (Sivaganesan et al., 2008) was used to generate standard curves. The lower limit of quantification (LLOQ) was defined as the instrument run specific 95 % credible interval upper-bound from repeated measures ($n = 18$) of SRM 2917 Level 1. Outliers were defined as the absolute value of a studentized residual >3. Pearson product momentum correlation analysis was performed between bacterial and viral fecal indicator paired measurements ($\alpha = 0.05$). A one-way analysis of variance (ANOVA) with Tukey's multiple comparisons tests was used to assess the differences in concentrations of *E. coli*, enterococci, and somatic coliphage across the sampling sites at a significance level of

Table 1
Host-associated qPCR primers and probes.

Assay	Reported target	Primer and probe sequences (5′ to 3′)	Reference
HF183/BacR287	Human	Forward: ATCATGAGTTCACATGTCCG Reverse: CTTCCTCTCAGAACCCCTATCC Probe: [FAM]-CTAATGGAACGCATCCC-[MGB]	(Green et al., 2014b; Bernhard and Field, 2000)
Rum2Bac	Ruminant	Forward: ACAGCCCGCGATTGATACTGGTAA Reverse: CAATCGGAGTTCTTCGTGAT Probe: [FAM]- ATGAGGTGGATGGAATTCGTGGTGT-[BHQ-1]	(Mieszkin et al., 2010)
DG3	Dog	Forward: TTTTCAGCCCCGTTGTTTCG Reverse: TGAGCGGGCATGGTCATATT Probe: [FAM]-AGTCTACGCGGGCGTACT-[MGB]	(Green et al., 2014a)
GFD	Avian	Forward: TCGGCTGAGCACTCTAGGG Reverse: GCGTCTCTTTGTACATCCCATTG Probe: [FAM/ZEN]-ACGTCAAGTCATCATGGCCCTTACGC-[IBFQ]	(Green et al., 2012; Weller et al., 2020)

FAM, 6-carboxyfluorescein; BHQ-1, black hole quencher 1; ZEN quencher; IBFQ, Iowa Black FQ.

$\alpha = 0.05$ (GraphPad Software Version 8.3.1, La Jolla, CA, USA). Due to the anticipated high frequency of non-detections (ND, $C_q = 40$) in host-associated genetic marker datasets, a Bayesian-based censored data analysis approach (Diedrich et al., 2023) (Fecal Score) was used to compare sites by fecal pollution source and to investigate potential links between host-associated genetic marker measurements, rainfall, and fecal indicator levels.

Fecal Score estimates represent a weighted average concentration [$\log_{10}$ copies per 100 mL with 95 % Bayesian credible interval (BCI)] calculated from a defined group of samples using all replicate measurements including non-detections (ND), detections below the LLOQ (BD), and measurements within the range of quantification (ROQ). Briefly, each replicate C_q measurement of a given water sample, belongs to one of the following three groups: ND ($C_q = 40$), BD (LLOQ $< C_q <$ 40), or ROQ ($C_q \leq$ LLOQ). The number of copies Z_1 for the ND group is assumed to be less than one. A Poisson distribution with mean λ_1 ($\lambda_1 < 1$) is assumed for Z_1. The number of copies Z_2 for the BD group is assumed to be between one and the number of copies (UB) corresponding to the LLOQ. A Poisson distribution with mean parameter λ_2 ($1 \leq \lambda_2 \leq$ UB) was assumed for Z_2. For the ROQ group, a 'mixed' calibration model (Sivaganesan et al., 2010) was used to estimate the number of copies Z_3 (in $\log_{10}$ base). For a given sample Y_j from the i^{th} instrument run, a weighted average of the posterior distributions of Z_1, Z_2 and Z_3 were used to estimate the number of copies X_j (in $\log_{10}$ base). The Bayesian models used for Z_1, Z_2 and Z_3 and the weighted average are given below:

For ND group:
$Z_1 \sim \text{Poisson}\ (\lambda_2) \cdot I\ (0, 1)$
$\log_{10}(\lambda_1) \sim N\ (0, 10^3) \cdot I\ (-2, 0).$
For BD group,
$Z_2 \sim \text{Poisson}\ (\lambda_1) \cdot I\ (1, UB_i)$
$UB_i = 10^{\frac{LLOQ_i - \alpha_i}{\beta}}$
$\log_{10}(\lambda_2) \sim N\ (0, 10^3) \cdot I\ (0,).$
For ROQ group:

$$log_{10}(Z_3) = \begin{cases} (C_{q1} - \alpha_i)/\overline{\beta} & \text{if}\, r_3 = 1 \\ (C_{qr} - \alpha_i)/\overline{\beta} & \text{if}\, r_3 > 1 \end{cases} \quad (1)$$

$$C_{qr} \sim N(\overline{C}_q, s^2).$$

$$\text{Log}_{10}X_j = [r_1 \cdot log_{10}(Z_1) + r_2 log_{10}(Z_2) + r_3 \cdot log_{10}(Z_3)]/(r_1 + r_2 + r_3)$$

where r_1, r_2 and r_3 are the number of C_q replicate measurements in ND, BD, ROQ groups, respectively, α_i is the intercept parameter for i^{th} instrument run and $\overline{\beta}$ is the master slope parameter of the mixed model calibration curve, $\overline{C}_q$ is the mean and s is the standard deviation of $\overline{C}_q$ derived from all measurements in the ROQ group. Moreover, I (a, b) restricts the Poisson and Normal distributions to a range from a to b and I (0,) restricts these distributions to positive values. For given samples Y_1, Y_2, …,Y_m, the posterior distribution of the Fecal Score (in $\log_{10}$ base) can be used to estimate the mean concentration and the corresponding 95 % BCI, where

$$\text{Log}_{10}(\text{Fecal Score}) = \sum_{1}^{m} log_{10}(X_j)/m. \quad (2)$$

A Fecal Score ratio was used to test for a significant difference between two Fecal Scores and is defined as the difference in Log_{10} (Fecal Score) between two sub-groups. If the Fecal Score ratio 95 % BCI does not include zero ($\log_{10}$ ratio of 1), then the two corresponding datasets are assumed to be significantly different.

A sample group was eligible for Fecal Score determination if it consisted of 1) at least five samples with one or more replicates classified as BD or ROQ, and 2) at least one sample with one or more replicate ROQ measurements. To rank sites, Fecal Scores were calculated by site for each eligible MST genetic marker data set using all samples. To investigate the potential influence of *E. coli*, enterococci, somatic coliphage, F+ coliphage, and rainfall levels on fecal pollution trends, water samples were organized by site and across sites into two groups prior to Fecal Score determination. For *E. coli*, groups included: 1) samples with *E. coli* < 235 CFU/100 mL and 2) samples with *E. coli* ≥ 235 CFU/100 mL [Beach Action Value (BAV) for estimated illness rate of 36 per 1000 primary contact recreators] (United States Environmental Protection Agency, 2012). For enterococci, groups included: 1) samples with enterococci <70 CFU/100 mL and 2) samples with enterococci ≥70 CFU/100 mL (BAV for estimated illness rate of 36 per 1000 primary contact recreators) (United States Environmental Protection Agency, 2012). There are currently no recommended BAV available for coliphage (somatic and F+). Instead, predicted risk-based thresholds were employed based on a study using quantitative microbial risk assessment (Boehm, 2019). For somatic and F+ coliphage, groups consisted of 1) samples with somatic <600 PFU/L or F+ < 300 PFU/L, and 2) samples with somatic ≥600 PFU/L and F+ ≥ 300 PFU/L. For rainfall, groups included: 1) 48 h cumulative precipitation = 0 mm, and 2) 48 h cumulative precipitation >0 mm. All statistics were conducted with SAS software (version 9.48 M8, Cary, NC) and WinBugs (version 1.4.3; https://www.mrc-bsu.cam.ac.uk/software/bugs/).

3. Results

3.1. Precipitation

Rainfall events were recorded using a USGS measurement station (Fig. 1). During the four-month study, cumulative 48 h precipitation measurements prior to sampling ranged from 0 mm to 26 mm with 43 % of sampling events (*n* = 18) defined as 'wet' conditions.

3.2. Culturable viral and bacterial fecal indicators

Somatic coliphage were present in all samples ranging from 0.3 to

4.17 $\log_{10}$ PFU/L (Fig. 2). Somatic coliphage levels exceeded a previously reported estimated risk-based threshold (600 PFU/L) in 36 % of samples, ranging from 7 % (HB) to 71 % (WP) by site. Across sites somatic coliphage levels were found to significantly correlate with both *E. coli* ($P < 0.0001$, $R^2 = 0.667$) and enterococci ($P = 0.0002$, $R^2 = 0.551$) paired measurements. Between site comparisons revealed that at site HB, significantly lower concentrations of both *E. coli* (F (2, 37) = 8.018, P ≤ value range: 0.0014 to 0.0162) and somatic coliphage (F (2, 37) = 22.43, P < ≤0.0001) were observed compared to the other sites.

Enterococci and *E. coli* concentrations are shown in Fig. 2. *E. coli* were present in all samples ranging from 1.60 to 3.93 $\log_{10}$ CFU/100 mL across sites. A total of 29 samples (70.7 %) exceeded the BAV of 235 CFU/100 mL for *E. coli* (United States Environmental Protection Agency, 1986). 'Wet' sampling corresponded with *E. coli* exceedances at both FC and WP catchments where 85 % of 'wet' samples resulted in *E. coli* measurements exceeding 235 CFU/100 mL. In contrast, only 36 % of *E. coli* exceedances occurred under 'wet' sampling conditions at HB. For enterococci, culturable levels were present in all samples with concentrations ranging from 2.19 to 3.95 $\log_{10}$ CFU/100 mL. Enterococci levels exceeded the water quality standard (70 CFU/100 mL) in all samples regardless of precipitation level. Overall, enterococci concentrations were higher than those observed for *E. coli* (Stocker et al., 2018; Litsky et al., 1955) but still exhibited a significant correlation with paired *E. coli* measurements ($P < 0.0001$, $R^2 = 0.739$).

3.3. Host-associated qPCR quality controls and measurements

To identify high-quality MST qPCR data, a series of quality controls were employed. Calibration models exhibited amplification efficiencies ranging from 0.97 (HF183/BacR287) to 1.00 (GFD). A total of 13 outliers (3.6 %) were observed (360 total measurements) with 84.6 % corresponding to the lowest standard concentration used as template (SRM 2917 Level 1). Please refer to Supplemental Information (Table S2) for detailed calibration model parameters and LLOQ values. Extraneous DNA control reactions indicated no evidence of contamination ($n = 444$ reactions). Amplification inhibition was not identified in any duplex IAC HF183/BacR287 experiments (inhibition threshold range: 33.23 C_q to 34.11 C_q) and all samples passed SPC screening. Seventeen samples (41.5 %) required C_q adjustments (range: 0.11 to 2.50 C_q; median: 0.44 C_q). Instrument run-specific IAC proficiency testing yielded a 100 % pass rate with HF183/BacR287 NTC VIC C_q standard deviations ranging from 0.14 C_q to 0.19 C_q. The SPC proficiency test used to monitor for suitable DNA recovery for each extraction batch indicated a 100 % pass rate with Sketa22 method blank standard deviations ranging from 0.03 C_q to 0.55 C_q.

Host-associated genetic markers were quantified for eligible water samples ($n = 41$) using MST qPCR methods for human (HF183/BacR287), dog (DG3), ruminant (Rum2Bac), and avian (GFD) pollution sources. Estimated mean $\log_{10}$ copies per reaction concentrations are shown in Fig. 3. The number of samples where all replicate reactions yielded a 40 C_q (non-detection) ranged from 18 (43.9 %, HF183/BacR287) to 38 (92.7 %, DG3).

3.4. Host-associated genetic marker fecal score trend analyses

A Fecal Score approach was used to compare sites and to evaluate the potential links between coliphage (somatic and F+), FIB (*E. coli* and enterococci), and rainfall based on a particular pollution source. To compare sites based on human (HF183/BacR287) and ruminant (Rum2Bac) fecal sources, a Fecal Score was estimated for each eligible qPCR assay and site combination (Table 2). Of the six assay and site combinations (2 qPCR assays · 3 sites = 6 combinations), 83.3 % were eligible for Fecal Score estimation. Rum2Bac site Fecal Scores suggest the ruminant fecal pollutions levels are similar (95 % BCI overlap), on

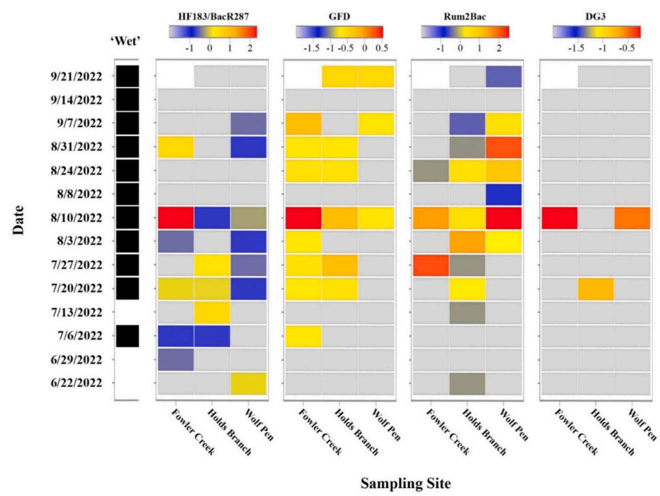


Fig. 3. Host-associated genetic marker estimated $\log_{10}$ copies per reaction concentrations for HF183/BacR287 (human), GFD (avian), Rum2Bac (ruminant), and DG3 (dog). Blocks shaded black under 'Wet' column represent samples occurring during rain events (48 h cumulative precipitation ≥ 0.0 mm). Heatmap keys are shown for each qPCR assay data set indicating $\log_{10}$ copies per reaction color coding information. Cells shaded light grey indicate samples where all replicate reactions yielded a $C_q = 40$ (non-detection) and cells shaded dark grey represent samples yielding detections below the lower limit of quantification. Sampling date and site combinations denoted by a white cell represent samples that failed the sample processing control and were discarded from the study.

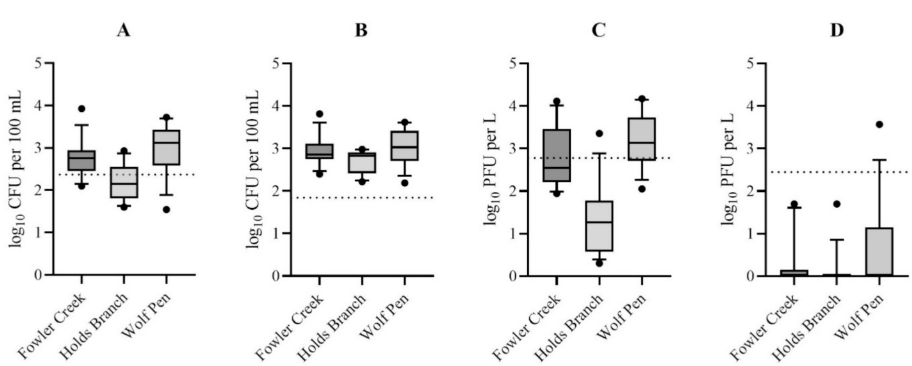


Fig. 2. Box and whisker plot for *E. coli* $\log_{10}$ CFU/100 mL (Panel A), enterococci $\log_{10}$ CFU/100 mL (Panel B), somatic coliphage $\log_{10}$ PFU/L (Panel C) and F+ coliphage $\log_{10}$ PFU/L (Panel D). Each box is delimited by 25th and 75th percentiles, the solid line within each box represents the median, and whiskers denote the 10th and 90th percentile values. Outliers are depicted as black dots. Horizontal dotted lines represent thresholds for *E. coli* (235 CFU/100 mL), enterococci (70 CFU/100 mL), somatic (600 PFU/L) and F+ coliphage (300 PFU/L).

Table 2
Fecal scores (log_{10} copies per 100 mL with 95 % BCI) for site comparisons with each eligible host-associated genetic marker.

Assay	Site	Fecal score	95 % BCI	
			Lower bound	Upper bound
HF183/BacR287	Wolf Pen	.	.	.
	Holds Branch	1.00	0.73	1.26
	Fowler Creek	1.23	0.98	1.49
Rum2Bac	Fowler Creek	1.18	0.91	1.45
	Holds Branch	1.44	1.22	1.65
	Wolf Pen	1.56	1.33	1.79

'.' indicates not eligible for Fecal Score determination.
BCI represents Bayesian Credible Interval.

average. Human-associated genetic marker datasets indicated a similar trend compared to Rum2Bac. Since F+ coliphage were only sporadically detected (detection frequency = 37 %; minimum = 1 PFU/100 mL; maximum = 74 PFU/100 mL) and all enterococci measurements exceeded the 70 CFU/100 mL BAV (lowest concentration = 133 CFU/100 mL), due to this, these datasets were excluded from statistical analysis.

To evaluate potential influences of coliphage, FIB, and rainfall, Fecal Scores were generated by categorizing qPCR data sets (across sites) into groups based on predefined threshold definitions: somatic coliphage (threshold: ≥ 600 PFU/100 mL), F+ coliphage (threshold: ≥ 300 PFU/100 mL), *E. coli* (threshold: ≥ 235 CFU/100 mL), enterococci (threshold: ≥ 70 CFU/100 mL), and rainfall (threshold: > 0 mm accumulation over 48 h prior to sampling). Because all samples were above the enterococci threshold and below the F+ threshold, Fecal Score determination for these definitions were excluded. When applying threshold definitions for somatic coliphage, *E. coli*, and rainfall across sites, 75 % of combinations were eligible to report Fecal Scores (2 qPCR assays · 3 threshold definitions · 2 groups/threshold definition = 12 possible combinations). The evaluation of Fecal Scores using threshold definitions across sites identified several host-associated genetic marker trends (Fig. 4). When samples were grouped based on somatic coliphage levels (threshold: ≥ 600 PFU/L), a significant difference (Fecal Score ratio = 0.524 log_{10} copies/100 mL, 95 % BCI lower bound = 0.243, upper bound = 0.807; 95 % BCI interval does not intersect zero) was observed for Rum2Bac Fecal Scores (Fig. 4, Panel A). A similar trend was identified when samples were grouped based on *E. coli* levels (threshold: ≥ 235 CFU/100 mL) where a significant difference (Fecal Score ratio = 0.550 log_{10} copies/100 mL, 95 % BCI range: 0.216 to 0.881; 95 % BCI interval does not intersect zero) occurred with Rum2Bac (Fig. 4, Panel B). In contrast, HF183/BacR287 across-site Fecal Scores were significantly different (Fecal Score ratio = 0.356 log_{10} copies/100 mL, 95 % BCI range: 0.050 to 0.664; 95 % BCI interval does not intersect zero) when samples were grouped using the somatic coliphage threshold (Fig. 4, Panel A). However, samples grouped using the *E. coli* threshold were not eligible for Fecal Score determination when below the local single sample maximum but did yield a Fecal Score estimate for samples with *E. coli* levels ≥235 CFU/100 mL (Fig. 4, Panel B). A stark difference was also observed when samples were grouped by the presence or absence of rainfall where both HF183/BacR287 and Rum2Bac were ineligible for Fecal Score determination in the absence of precipitation because no replicates yielded measurements in the ROQ, but host-

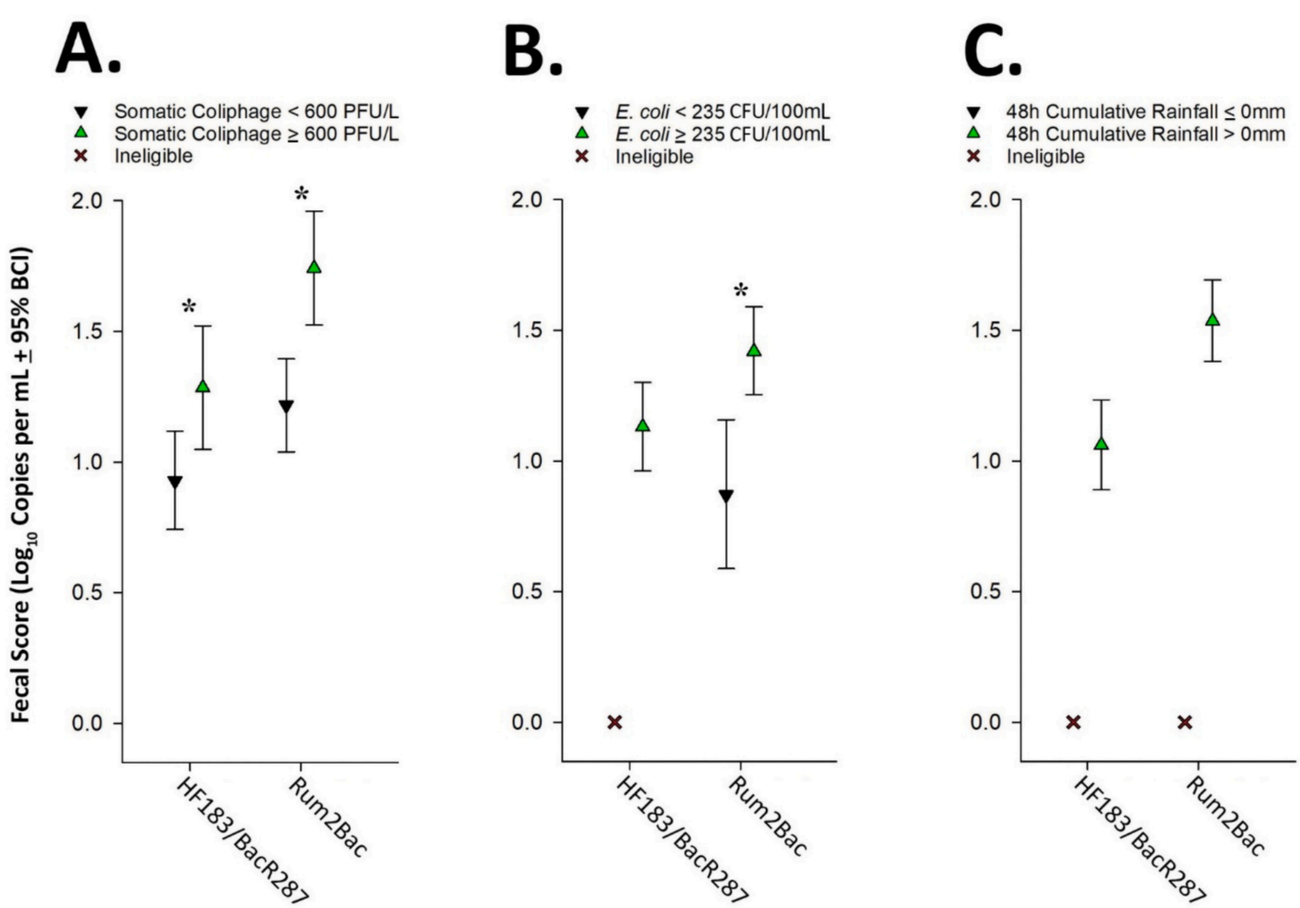


Fig. 4. Host-associated genetic marker Fecal Scores (log_{10} copies per 100 mL with 95 % BCI) using somatic coliphage (Panel A; threshold: ≥600 PFU/L), *E. coli* (Panel B; threshold: ≥235 CFU/100 mL), and 48 h rain accumulation prior to sampling (Panel C; threshold: >0 mm) sample group definitions. An '*' denotes significant difference (Fecal Score ratio 95 % BCI do not intersect zero) between Fecal Scores. An 'X' denotes that a sample grouping was ineligible for Fecal Score determination.

associated genetic markers indicated quantifiable levels in samples collected after rain events (Fig. 4, Panel C). When samples were grouped by somatic coliphage, *E. coli*, and rainfall definitions by site, the majority (72.2 %) of possible assay, site, and definition combinations (2 qPCR assays · 3 sites · 3 threshold definitions · 2 groups/threshold definition = 36 possible combinations) were ineligible for Fecal Score determination (data not shown). The most notable trend from by site Fecal Score analyses was the ineligibility of all qPCR assays and site combinations in the absence of rainfall.

4. Discussion

4.1. Bacterial and viral fecal indicator trends

General fecal indicator *E. coli* and enterococci water quality monitoring in small stream tributaries identified multiple fecal pollution insights. First, FIB measurements indicate considerable fecal contamination where 70.7 % of samples exceeded the *E. coli* BAV (235 CFU/100 mL) and 100 % of samples exceeded the enterococci BAV (70 CFU/100 mL), suggesting that small stream tributaries can harbor high concentrations of fecal waste that can potentially compromise water quality within larger receiving waterbodies. Second, while *E. coli* and enterococci paired measurements were significantly correlated ($P < 0.0001$, $R^2 = 0.739$), enterococci consistently occurred at higher levels. Elevated enterococci levels are documented in other flowing freshwater studies (Matthew Daniel Stocker et al., 2019; Browning et al., 2023; Jeon et al., 2020) and, in part, may be the result of naturalized populations (Ferguson et al., 2005; Piggot et al., 2012) and/or natural reservoirs such as submerged aquatic vegetation (Badgley et al., 2010a, 2010b). Additional research is needed to identify sources of enterococci in Banklick Creek watershed small stream tributaries. Third, somatic coliphage measurements exceeded a proposed advisory threshold level (600 PFU/L) in 34.1 % of sampling events, markedly lower than FIB (*E. coli* and enterococci). Together, these findings suggest that somatic coliphage water quality assessment in low-order streams may be a plausible alternative; however additional research is needed to determine whether FIB or somatic coliphage are more closely associated with pathogenic microbes. F+ coliphage exhibited a different trend where detection frequency was low (< 25 %) with no sample exceeding the suggested advisory of 300 PFU/L. Low F+ detection frequencies have also been reported at fresh (Cyterski et al., 2022; McMinn et al., 2017a) and marine (McMinn et al., 2017b; Rodriguez et al., 2012) recreational beaches and could be attributed to differential host shedding patterns. For example, F+ coliphages are often shed at lower concentrations (median 4.5 $\log_{10}$ PFU/L) in human waste compared to *E. coli* (median 6.4 $\log_{10}$ CFU/100 mL), enterococci (median 5.6 $\log_{10}$ CFU/100 mL), and somatic coliphage (median 4.9 $\log_{10}$ PFU/L) (Korajkic et al., 2022; McMinn et al., 2017a). In addition, F+ can originate from numerous non-human sources, albeit usually at much lower levels compared to human waste (McMinn et al., 2014). These findings suggest that a 1 L sample volumes may not be sufficient for routine F+ coliphage for surface water quality monitoring. Finally, FIB measurements demonstrate that fecal pollution is pervasive regardless of Scenario (1, 2, or 3) in the study areas, suggesting that catchments with markedly different land use activities and variable waste management practices can all contribute to poor water quality.

4.2. Host-associated genetic marker occurrence trends

Systematic measurement of host-associated genetic markers in small stream tributary samples revealed multiple trends. For example, the dog-associated genetic marker (DG3) was rarely detected at any site, suggesting that dog waste does not play an important role in water quality exceedances in the study area. While this finding is not surprising for Scenario 3 (WP catchment; agricultural and forest covered), other studies report the presence of dog waste routinely in receiving and storm waters situated in catchments with a large residential component (Diedrich et al., 2023; Staley et al., 2016; Ervin et al., 2014), suggesting either low concentrations of canines or local authorities are successfully managing pet waste in these areas.

Although the avian-associated marker (GFD) was regularly detected in each catchment, it typically did not occur at quantifiable levels, suggesting minimal presence of avian wildlife over the study period. In contrast, a previous study focusing on a constructed wetland to mitigate poor water quality within the Banklick Creek concluded there was a considerable impact from avian activities on wetland treatment efficacy as wetlands offer an ideal habitat, encouraging waterfowl congregation (i.e., open body of water next to nesting sites) (McMinn et al., 2019). However, a similar trend was not observed in select Banklick Creek small stream tributaries in this study suggesting that birds may not frequent small streams with enough regularity to impact water quality and/or bird population sizes shift at different times of year. Additional research is needed to describe bird behavior in small stream tributary scenarios, to document local bird migration activities, and to ascertain if avian waste could have an impact on water quality at other times of the year.

Contrary to dog and avian sources, human and ruminant-associated genetic markers were observed on a more consistent basis and at higher concentrations. Though the HF183/BacR287 human-associated genetic marker was present at all sites, detection frequencies were highest at WP (Scenario 3) and FC (Scenario 2) catchments, 50 % and 46 %, respectively (Fig. 3). The elevated occurrence of the human-associated marker in these catchments could be a result of the high septic system counts in these catchments (WP = 70; FC = 444). Septic systems require routine maintenance and if ignored, system failures could be a source of human fecal contamination, especially after rain events. For the ruminant-associated marker (Rum2Bac), detection frequencies were most prevalent at HB (Scenario 1) and WP (Scenario 3) catchments, 64 % and 50 %, respectively. The HB catchment contained the highest concentration of residential parcels, impervious surfaces, and stormwater outfalls. A recent study reports that stormwater outfall discharges from highly urbanized catchments can contain ruminant fecal pollution (Shanks et al., 2024), suggesting a situation where residential catchments consisting of lawns and shrubs may attract wildlife, such as foraging white-tailed deer (Nagy-Reis et al., 2019; Darlington et al., 2022), resulting in the accumulation of scat overtime that are subsequently washed down storm drains and discharged into small streams during rain events. For the WP catchment (Scenario 3), high detection rates of Rum2Bac may be attributed to the presence of both wildlife and small-scale cattle operations in the area, with some cattle having direct access to tributaries, suggesting that the addition of fencing and or buffer zones denying these animals direct access could improve water quality. Additionally, the WP catchment also contained the highest agricultural land cover (3.04 km^2), where it is common to use livestock waste for soil fertilization to promote crop growth (Newton et al., 2003; Savage and Ribaudo, 2013). Systematic testing of small stream tributaries representing three different land use scenarios demonstrated the advantage of combining host-associated genetic marker testing over relying solely on general fecal indicators for water quality management.

4.3. Use of the fecal score approach to explore potential fecal indicator, precipitation, and host-associated genetic marker trends

Due to the high frequency of host-associated genetic marker measurements yielding detections below the LLOQ (BD) or non-detections (ND), a Bayesian-based censored data analysis approach was used, allowing for the incorporation of all qPCR-based measurements from a defined group of samples. Fecal score analyses revealed multiple links between somatic coliphage, *E. coli*, precipitation, and fecal pollution sources. First, across-site Fecal Scores were significantly higher (Fecal Score ratio 95 % BCI does not intersect zero) when *E. coli* exceeded the BAV ($\geq$ 235 CFU/100 mL) or somatic coliphage was >600 PFU/L indicating that sources such as cattle and local deer populations likely contribute to poor water quality. Second, the human-associated genetic marker (HF183/BacR287) exhibited a similar trend, but at different magnitudes depending on whether water samples were grouped based on *E. coli* (Fig. 4, Panel A) or a somatic coliphage-based single sample maximum definition (Fig. 4, Panel B), suggesting that somatic coliphage originating from human sources may be more susceptible to decay compared to cultivated *E. coli* once discharged into the small tributary stream environment. This trend parallels findings from a persistence study reporting that the attenuation of somatic coliphage and cultivated *E. coli* varied by source of human fecal pollution (i.e., septage compared to untreated sewage and human fecal material) (Wanjugi et al., 2016), suggesting that septic pollution may be a key contributor of human fecal pollution in Banklick small stream water quality. Additional research is needed to investigate further the influence of different human fecal pollution sources on the occurrence of somatic coliphage compared to cultivated *E. coli* and the specific sources of human fecal contamination in Banklick small stream tributaries (septic vs sewage). Finally, across-site Fecal Scores derived from samples categorized into groups based on the presence or absence of rainfall identified precipitation as a key mechanism for the introduction of both human and ruminant waste into Banklick tributaries. In the absence of rain, both HF183/BacR287 and Rum2Bac host-associated genetic markers were negligible to the extent that Fecal Score determination was not possible for these datasets. This suggests a source loading scenario in which ruminant fecal pollution deposited on soil surfaces and human waste introduced via failing waste treatment systems are washed into tributaries during precipitation events. The impact of rain events on water quality is further supported by the observation that no samples collected in the absence of rainfall exceeded the *E. coli* BAV. While many studies report the importance of rainfall for surface water quality (Dwight et al., 2011; Zhao et al., 2018; Kohzu et al., 2023), results presented here clearly demonstrate the strong link between water quality and precipitation in small stream tributary systems.

4.4. Factors to consider for fecal score implementation

MST data sets often consist of a large proportion of non-detections or measurements below the LLOQ (Li et al., 2021, 2019; Diedrich et al., 2023), resulting in a 'censored' data situation where the target concentration cannot be firmly established for these samples. 'Censored' data sets require statistical approaches that can responsibly incorporate these measurements into quantitative findings (Cao et al., 2018). To date, the Fecal Score approach has been used to prioritize sites for remediation (Li et al., 2019), investigate the impact of rainfall at recreational beaches (Diedrich et al., 2023), and study the influence of municipal storm sewer system discharges on urban stream water quality (Shrestha et al., 2020). While this 'censored' data method uncovered several important trends regarding fecal indicator monitoring and influence of rainfall on small stream tributaries in this study, many dataset groupings were deemed ineligible for Fecal Score determination for site comparisons, with multiple factors likely contributing to this outcome. First, findings indicate that small stream water quality within the Banklick watershed is highly responsive to rain events, with water quality substantial better under baseflow conditions, leading to a high number of non-detections (ND) in 'dry' samples (57 % of total samples). Secondly, two fecal pollution sources targeted in this study appear to have a minimal impact on water quality at study sites, where DG3 (dog) and GFD (avian) were virtually undetectable or occurred at very low concentrations, respectively. These findings suggest that when a particular fecal pollution source consistently occurs at a low frequency and/or concentration, resulting host-genetic marker data sets will not be eligible for Fecal Score determination regardless of the number of sampling events. And lastly, the total number of samples collected from each site proved to be inadequate (14 samples per site) to consistently evaluate the influence of fecal indicator monitoring approaches and rainfall on host-associated genetic marker occurrence on a site basis. Especially when evaluating the influence of a specific factor (e.g., rainfall) at a given site. This is because the number of samples from a site must be categorized into two groups (e.g., group 1 = absence of rainfall, group 2 = presence of rainfall), limiting the number of samples available to calculate a Fecal Score. Together, and for future applications, the above observations suggest that water quality in small stream tributary systems like those within the Banklick watershed may require a larger number of sampling events and/or replicate measurements per sample to maximize the Fecal Score data analysis approach. Findings also indicate that conducting pilot testing prior to formal investigations could be a useful strategy to tailor field sampling strategies and host-associated genetic marker selection.

In conclusion, this study provides new insights into small stream tributary monitoring with FIB, viral indicators, host-associated genetic markers, and land use information. Key findings demonstrate that some viral indicators may be more suitable than others for small stream tributary monitoring as well as highlight the potential impact of catchment land use activities and human fecal pollution source (e.g., septic compared to sewage) on indicator occurrence. Results also underscore the important role of precipitation in small stream tributary water quality and the strengths and limitations of the Fecal Score censored data analysis approach to quantify fecal pollution source associated impacts on water quality.

CRediT authorship contribution statement

Brian R. McMinn: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Asja Korajkic:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Julie Kelleher:** Writing – original draft, Methodology, Investigation, Formal analysis. **Adam Diedrich:** Writing – original draft, Methodology, Investigation, Formal analysis. **Adin Pemberton:** Writing – original draft, Methodology, Investigation, Formal analysis. **Jessica R. Willis:** Writing – original draft, Methodology, Investigation, Formal analysis. **Mano Sivaganesan:** Writing – original draft, Software, Formal analysis, Data curation. **Brooke Shireman:** Writing – original draft, Methodology, Investigation. **Andrew Doyle:** Writing – original draft, Methodology, Investigation. **Orin C. Shanks:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This study was conducted in partnership with Sanitation District 1 (SD1, Kentucky, USA). The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency or SD1. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.175740.

References

Badgley, B.D., Thomas, F.I., Harwood, V.J., 2010a. The effects of submerged aquatic vegetation on the persistence of environmental populations of Enterococcus spp. Environ. Microbiol. 12 (5), 1271–1281.

Badgley, B.D., Nayak, B.S., Harwood, V.J., 2010b. The importance of sediment and submerged aquatic vegetation as potential habitats for persistent strains of enterococci in a subtropical watershed. Water Res. 44 (20), 5857–5866.

Baral, D., et al., 2018. Quantifying the relative contributions of environmental sources to the microbial community in an urban stream under dry and wet weather conditions. Appl. Environ. Microbiol. 84 (15).

Barnes, A.N., et al., 2018. The association between domestic animal presence and ownership and household drinking water contamination among peri-urban communities of Kisumu, Kenya. PLoS One 13 (6), e0197587.

Bergquist, A.M., et al., 2016. Evaluation of a hybrid ion exchange-catalyst treatment technology for nitrate removal from drinking water. Water Res. 96, 177–187.

Bernhard, A.E., Field, K.G., 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA. Appl. Environ. Microbiol. 66 (10), 4571–4574.

Bianchi, E., et al., 1995. Comparison between ion chromatography and a spectrophotometric method for determination of nitrates in meat products. Z. Lebensm. Unters. Forsch. 200 (4), 256–260.

Boehm, A.B., 2019. Risk-based water quality thresholds for coliphages in surface waters: effect of temperature and contamination aging. Environ Sci Process Impacts 21 (12), 2031–2041.

Brooks, Y.M., et al., 2020. Fecal indicator bacteria, fecal source tracking markers, and pathogens detected in two Hudson River tributaries. Water Res. 171, 115342.

Browning, Dustin A., W.E.M., Stookey, Courtney, Nikolai, Stephen J., Barrow, Joel, Townsend II, Darrell E., 2023. Validating microbial source tracking markers and assessing the efficacy of culturable *E. coli* and Enterococcus assays in Ozark streams, USA. Water Air Soil Pollut. 234.

Bustin, Stephen A., 2006. International University Line, AZ of Quantitative PCR (La Jolla, California).

Cao, Y., et al., 2018. A human fecal contamination score for ranking recreational sites using the HF183/BacR287 quantitative real-time PCR method. Water Res. 128, 148–156.

Cyterski, M., et al., 2022. Bacterial and viral fecal indicator predictive modeling at three Great Lakes recreational beach sites. Water Res. 223, 118970.

Darlington, S., et al., 2022. Cumulative effects of human footprint, natural features and predation risk best predict seasonal resource selection by white-tailed deer. Sci. Rep. 12 (1), 1072.

Dean, K., Mitchell, J., 2022. Identifying water quality and environmental factors that influence indicator and pathogen decay in natural surface waters. Water Res. 211, 118051.

Diedrich, A., et al., 2023. Genetic fecal source identification in urban streams impacted by municipal separate storm sewer system discharges. PLoS One 18 (1), e0278548.

Dwight, R.H., et al., 2011. Influence of variable precipitation on coastal water quality in southern California. Water Environ. Res. 83 (12), 2121–2130.

Ervin, J.S., et al., 2014. Microbial source tracking in a coastal California watershed reveals canines as controllable sources of fecal contamination. Environ. Sci. Technol. 48 (16), 9043–9052.

Ferguson, D.M., et al., 2005. Enumeration and speciation of enterococci found in marine and intertidal sediments and coastal water in southern California. J. Appl. Microbiol. 99 (3), 598–608.

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. 41 (16), 3517–3538.

Garfield, L., Walker, M., 2008. Microbial water quality and influences of fecal accumulation from a dog exercise area. J. Environ. Health 71 (4), 24–29.

Gitter, A., et al., 2023. Integrating microbial source tracking with quantitative microbial risk assessment to evaluate site specific risk based thresholds at two South Florida beaches. Front. Microbiol. 14, 1210192.

Green, H.C., et al., 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. Appl. Environ. Microbiol. 78 (2), 503–510.

Green, H.C., et al., 2014a. Development of rapid canine fecal source identification PCR-based assays. Environ. Sci. Technol. 48 (19), 11453–11461.

Green, H.C., et al., 2014b. Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Appl. Environ. Microbiol. 80 (10), 3086–3094.

Green, H., et al., 2019. Microbial source-tracking reveals origins of fecal contamination in a recovering watershed. Water (Basel) 11 (10).

Harwood, V.J., et al., 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. FEMS Microbiol. Rev. 38 (1), 1–40.

Haugland, R.A., et al., 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected Bacteroidales species and human fecal waste by qPCR. Syst. Appl. Microbiol. 33 (6), 348–357.

Holcomb, D.A., Stewart, J.R., 2020. Microbial indicators of fecal pollution: recent progress and challenges in assessing water quality. Curr Environ Health Rep 7 (3), 311–324.

Jeon, D.J., et al., 2020. Temporal stability of E. coli and Enterococci concentrations in a Pennsylvania creek. Environ. Sci. Pollut. Res. Int. 27 (4), 4021–4031.

Kistemann, T., et al., 2002. Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. Appl. Environ. Microbiol. 68 (5), 2188–2197.

Kline, M.C., et al., 2009. Production and certification of NIST Standard Reference Material 2372 Human DNA Quantitation Standard. Anal. Bioanal. Chem. 394 (4), 1183–1192.

Kohzu, A., et al., 2023. Identifying the true drivers of abrupt changes in ecosystem state with a focus on time lags: extreme precipitation can determine water quality in shallow lakes. Sci. Total Environ. 881, 163097.

Korajkic, A., et al., 2022. Effectiveness of two wastewater disinfection strategies for the removal of fecal indicator bacteria, bacteriophage, and enteric viral pathogens concentrated using dead-end hollow fiber ultrafiltration (D-HFUF). Sci. Total Environ. 831, 154861.

Li, X., et al., 2019. Large-scale implementation of standardized quantitative real-time PCR fecal source identification procedures in the Tillamook Bay Watershed. PLoS One 14 (6), e0216827.

Li, X., et al., 2021. Variable fecal source prioritization in recreational waters routinely monitored with viral and bacterial general indicators. Water Res. 192, 116845.

LifeTechnologies, 2014. Real-time PCR Handbook, 3rd edition, p. 11.

Light Gray Canvas Base Map. In: ESRI Light Gray Canvas Base Map, 2023. Esri, TomTom, Garmin, FAO, NOAA, USGS, © OpenStreetMap contributors, and the GIS User Community.

Link-GIS Topo_2020_CC_KC Geospatial Data Presentation Form: vector digital data. In: CCKC Contours – 2020, 2021.

Litsky, W., Mallmann, W.L., Fifield, C.W., 1955. Comparison of the most probable numbers of Escherichia coli and enterococci in river waters. Am. J. Public Health Nations Health 45 (8), 1049–1053.

Lobos, A.E., Brandt, A.M., Gallard-Góngora, J.F., Korde, R., Brodrick, E., Harwood, V.J., 2024. Persistence of sewage-associated genetic markers in advanced and conventional treated recycled water: implications for microbial source tracking in surface waters. mBio, 0065524. https://doi.org/10.1128/mbio.00655-24.

Mallin, M.A., Johnson, V.L., Ensign, S.H., 2009. Comparative impacts of stormwater runoff on water quality of an urban, a suburban, and a rural stream. Environ. Monit. Assess. 159 (1–4), 475–491.

Matthew Daniel Stocker, J.E.S., Hernandez, Cesar, Macarisin, Dumitru, Pachepsky, Yakov, 2019. Seasonality of E. coli and enterococci concentrations in creek water, sediment, and periphyton. Water Air Soil Pollut. 230.

McMinn, B.R., Korajkic, A., Ashbolt, N.J., 2014. Evaluation of Bacteroides fragilis GB-124 bacteriophages as novel human-associated faecal indicators in the United States. Lett. Appl. Microbiol. 59 (1), 115–121.

McMinn, B.R., Ashbolt, N.J., Korajkic, A., 2017a. Bacteriophages as indicators of faecal pollution and enteric virus removal. Lett. Appl. Microbiol. 65 (1), 11–26.

McMinn, B.R., et al., 2017b. Concentration and quantification of somatic and F+ coliphages from recreational waters. J. Virol. Methods 249, 58–65.

McMinn, B.R., et al., 2019. A constructed wetland for treatment of an impacted waterway and the influence of native waterfowl on its perceived effectiveness. Ecol. Eng. 128, 48–56.

Mieszkin, S., et al., 2010. Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol. 108 (3), 974–984.

Nagy-Reis, M.B., et al., 2019. Conservation Reserve Program is a key element for managing white-tailed deer populations at multiple spatial scales. J. Environ. Manag. 248, 109299.

National Health and Medical Research Council, 2008. Austrailian Government, Guidelines for Managing Risks in Recreational Water. Commonwealth Copyright Administration: National Circuit, Canberra, ACT, ISBN 1864962666.

Newton, G.L., et al., 2003. Managing manure nutrients through multi-crop forage production. J. Dairy Sci. 86 (6), 2243–2252.

Piggot, A.M., et al., 2012. Relationship between enterococcal levels and sediment biofilms at recreational beaches in South Florida. Appl. Environ. Microbiol. 78 (17), 5973–5982.

Price, M.T., Blackwood, A.D., Noble, R.T., 2023. Cross-method assessment of coliphage detection methods in the framework of EPA regulatory standards. Lett. Appl. Microbiol. 76 (7).

Raith, M.R., et al., 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Res. 47 (18), 6921–6928.

Reynolds, L.J., et al., 2021. Identifying sources of faecal contamination in a small urban stream catchment: a multiparametric approach. Front. Microbiol. 12, 661954.

Rodriguez, R.A., et al., 2012. Comparison of methods for the detection of coliphages in recreational water at two California. United States beaches. J Virol Methods 181 (1), 73–79.

Sanitiation District No.1 of Northern Kentucky, 2023. Watershed characterization report. httpa://www.sd1.org/233/Watershed-Characterization-Reports.

Savage, J.A., Ribaudo, M.O., 2013. Impact of environmental policies on the adoption of manure management practices in the Chesapeake Bay watershed. J. Environ. Manag. 129, 143–148.

Shanks, O.C., et al., 2016. Data acceptance criteria for standardized human-associated fecal source identification quantitative real-time PCR methods. Appl. Environ. Microbiol. 82 (9), 2773–2782.

Shanks, O.C., et al., 2024. Quantitative fecal source characterization of urban municipal storm sewer system outfall ‘wet’ and ‘dry’ weather discharges. Water Res. 259, 121857.

Shrestha, A., et al., 2020. Fecal pollution source characterization at non-point source impacted beaches under dry and wet weather conditions. Water Res. 182, 116014.

Sivaganesan, M., et al., 2008. A Bayesian method for calculating real-time quantitative PCR calibration curves using absolute plasmid DNA standards. BMC Bioinformatics 9, 120.

Sivaganesan, M., et al., 2010. Improved strategies and optimization of calibration models for real-time PCR absolute quantification. Water Res. 44 (16), 4726–4735.

Staley, C., et al., 2012. Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. Appl. Environ. Microbiol. 78 (20), 7317–7326.

Staley, Z.R., et al., 2016. Comparison of microbial and chemical source tracking markers to identify fecal contamination sources in the Humber River (Toronto, Ontario, Canada) and associated storm water outfalls. Appl. Environ. Microbiol. 82 (21), 6357–6366.

Stocker, M.D., Penrose, M., Pachepsky, Y.A., 2018. Spatial patterns of Escherichia coli concentrations in sediment before and after high-flow events in a First-Order Creek. J. Environ. Qual. 47 (5), 958–966.

Systems, H, 2012. National Hydrography Dataset Plus - NHDPlus, in National Hydrography Dataset for Kentucky.

Tarek, M.H., Hubbart, J., Garner, E., 2023. Microbial source tracking to elucidate the impact of land-use and physiochemical water quality on fecal contamination in a mixed land-use watershed. Sci. Total Environ. 872, 162181.

United States Environmental Protection Agency, 1986. Office of Water, Quality Criteria for Water 1986 (Washington, D.C.). EPA 440/5-86-001.

United States Environmental Protection Agency, 1993. Office of Water. In: Method 300.1: Determination of Inorganic Anions in Drinking Water by Ion Chromatography (Washington, D.C.).

United States Environmental Protection Agency, 2009. Office of Water, Method 1600: Enterococi in Water by Membrane Filration Using mEI (Washington, D.C.). EPA-821-R-09-016.

United States Environmental Protection Agency, 2012. Office of Water, Recreational Water Quality Criteria (RWQC) for Bacterial Indicators of Fecal Contamination (Washington, D.C.). EPA - 820-F-12-061.

United States Environmental Protection Agency, 2014. Office of Water, Method 1603: *Escherichia coli* (*E. coli*) in Water by Membrane Filtration Using Modified Membrane-thermotolerant *Escherichia coli* Agar (Modified mTEC) (Washington, D.C.). EPA-821-R-09-007.

United States Environmental Protection Agency, 2015. Office of Water, Review of Coliphages as Possible Indicators of Fecal Contamination for Ambient Water Quality (Washington, D.C.). EPA-820-R-15-098.

United States Environmental Protection Agency, 2019. Office of Water, Method 1696: Characterization of Human Fecal Pollution in Water by HF183/BacR287 TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay (Washington, D.C.). EPA 821-R-19-002.

Vogel, J.R., et al., 2007. Identifying fecal sources in a selected catchment reach using multiple source-tracking tools. J. Environ. Qual. 36 (3), 718–729.

Wanjugi, P., et al., 2016. Differential decomposition of bacterial and viral fecal indicators in common human pollution types. Water Res. 105, 591–601.

Weller, D., et al., 2020. Landscape, water quality, and weather factors associated with an increased likelihood of foodborne pathogen contamination of New York streams used to source water for produce production. Front Sustain Food Syst 3.

Willis, J.R., et al., 2022. Performance of NIST SRM(R) 2917 with 13 recreational water quality monitoring qPCR assays. Water Res. 212, 118114.

Wittman, J., et al., 2013. Evaluation of land use and water quality in an agricultural watershed in the USA indicates multiple sources of bacterial impairment. Environ. Monit. Assess. 185 (12), 10395–10420.

Zhao, Z., et al., 2018. Studies on the spatiotemporal variability of river water quality and its relationships with soil and precipitation: a case study of the Mun River Basin in Thailand. Int. J. Environ. Res. Public Health 15 (11).

Zimmer-Faust, A.G., et al., 2023. Relationship between coliphage and Enterococcus at southern California beaches and implications for beach water quality management. Water Res. 230, 119383.